

Resolution of the stochastic strategy spatial prisoner's dilemma by means of particle swarm optimization

Jianlei Zhang,^{1,*} Chunyan Zhang,¹ Tianguang Chu,^{1,2,†} Matjaž Perc^{3,§}

1 State Key Laboratory for Turbulence and Complex Systems, College of Engineering, Peking University, Beijing, China **2** Key Laboratory of Machine Perception, Ministry of Education, Peking University, Beijing, China **3** Faculty of Natural Sciences and Mathematics, University of Maribor, Slovenia

*jianleizhang@pku.edu.cn

†chutg@pku.edu.cn

§matjaz.perc@uni-mb.si [www.matjazperc.com]

Abstract

We study the evolution of cooperation among selfish individuals in the stochastic strategy spatial prisoner's dilemma game. We equip players with the particle swarm optimization technique, and find that it may lead to highly cooperative states even if the temptations to defect are strong. The concept of particle swarm optimization was originally introduced within a simple model of social dynamics that can describe the formation of a swarm, i.e., analogous to a swarm of bees searching for a food source. Essentially, particle swarm optimization foresees changes in the velocity profile of each player, such that the best locations are targeted and eventually occupied. In our case, each player keeps track of the highest payoff attained within a local topological neighborhood and its individual highest payoff. Thus, players make use of their own memory that keeps score of the most profitable strategy in previous actions, as well as use of the knowledge gained by the swarm as a whole, to find the best available strategy for themselves and the society. Following extensive simulations of this setup, we find a significant increase in the level of cooperation for a wide range of parameters, and also a full resolution of the prisoner's dilemma. We also demonstrate extreme efficiency of the optimization algorithm when dealing with environments that strongly favor the proliferation of defection, which in turn suggests that swarming could be an important phenomenon by means of which cooperation can be sustained even under highly unfavorable conditions. We thus present an alternative way of understanding the evolution of cooperative behavior and its ubiquitous presence in nature, and we hope that this study will be inspirational for future efforts aimed in this direction.

Introduction

Cooperation is the basis for complex organizational structures in biological as well as social systems. Nevertheless, understanding the emergence and stability of cooperative behavior in the context of Darwinian selection remains a challenge to date. The dilemmas of cooperation are usually tackled within the framework of evolutionary game theory [1–3]. Although several mechanism allowing for the evolution of cooperation have already been identified [4], the resolution of social dilemmas and the closely related avoidance of the “tragedy of the commons” [5] is still considered an open problem. The prisoner's dilemma game [6], in particular, has attracted considerable attention in the past three decades [7–10], and to date it is widely consider as a paradigmatic example for the tensions between social welfare and individual interests [11–33]. Cooperation and defection are the two strategies that are at the heart of the prisoner's dilemma game. In general, while cooperators sacrifice some of their personal fitness for the benefit of the society, defectors succumb to the temptations and take full advantage of them. The prisoner's dilemma captures this situation by means of the following payoffs: mutual cooperation yields the reward R , mutual defection leads to punishment P , and the mixed choice gives the cooperator the sucker's payoff S and the defector the temptation T . The payoff ranking thus satisfies $T > R > P > S$. In the iterated prisoner's dilemma game the assumption that the mutual cooperation yields the highest collective income imposes another constraint, namely $2R > T + S$. This makes

it clear that the rational (selfish) action is to defect, and according to the fundamental principles of Darwinian selection, cooperation extinction is inevitable. Full defection is indeed the only stable Nash equilibrium for the prisoner's dilemma game in well-mixed populations.

Since the seminal paper by Nowak and May [34], however, we know that this may not be the case for spatial interactions. Although not universally applicable [35], spatial reciprocity is recognized as a potent promoter of cooperative behavior, even more so on complex networks [36–40] (for a comprehensive review see [8]). Other prominent mechanism promoting cooperation are kin selection [41], direct and indirect reciprocity [42–44], as well as group selection [45–47], to name but a few.

Inspired by previous works on this subject, we here introduce particle swarm optimization [48–50] to the players engaging in the prisoner's dilemma game on a square lattice [51], with the aim of investigating its impact on the evolution of cooperation. However, we abandon the commonly considered assumption that the players can choose only between the two pure strategies, namely to either cooperate or to defect. Real-life situations are often more complex than that, and indeed there is a lot of gray between the black and white extremes. Motivated by this fact, we here consider stochastic strategies, such that the cooperativeness of each players is determined by $W \in [0, 1]$. $W = 1$ returns full cooperation, while $W = 0$ returns full defection. These are the two extremes recovered from our present setup. Between $0 < W < 1$, however, there exists a continuous set of strategies that can be considered either as predominantly cooperative (if $W > 0.5$) or predominantly defective (if $W < 0.5$). Moreover, while the evolution of strategies is traditionally performed by means of different strategy adoption (or updating) rules (see [8] for a comprehensive review), we here take a much less explored avenue, namely by considering the aforementioned particle swarm optimization as the driving force behind strategy evolution. The particle swarm optimization algorithm is based on a simplified social model that is tightly tied to the theory of swarming [48–50]. A traditional analogy is a swarm of bees searching for a food source. In this analogy, each bee (considered here as a particle) makes use of its own memory as well as the knowledge obtained by the swarm as a whole, to find the best available food source. Particle swarm optimization can also be considered as being representative for multidimensional search (for example to find an optimum of a utility function). Typically, a number of simple entities (the “particles”) is randomly positioned in the search space, and to each a velocity vector is assigned, which is subsequently used to update the current position of each particle in the swarm. Each particle then proceeds by evaluating the objective function at its current location, and finally to determining its movement through the search space by combining some aspects of the history of its own current as well as other potentially optimal locations with those of one or more members of the swarm. Thus, the process makes use of the memory of each particle, as well as the knowledge gained by the swarm as a whole. The next iteration takes place after all the particles have moved once. Eventually the swarm, like a flock of birds collectively foraging for food, is likely to move closer to an optimum of the utility function. Accordingly, the particles (bees, birds, players) therefore should have a tendency to fly towards better and better areas over the course of the search process.

Here we focus specifically on introducing the particle swarm optimization algorithm to the strategy updating process in the stochastic strategy prisoner's dilemma game on the square lattice. In agreement with the above described general concept, each individual is assigned a variable from the unit interval determining its level of cooperativeness (or willingness to cooperate). Likewise, a velocity vector is assigned to every player. Following this initialization, each player makes use of its own memory (i.e., keeping score of the most profitable individual strategy in the past), as well as use of the knowledge gained by the swarm (i.e., the nearest neighbors) as a whole, to find the best available strategy for itself and the society. In particular, the particle swarm optimization algorithm makes use of the velocity vector to update the current strategy of each player in the swarm. In this sense our study can be considered related to previous works investigating the effects of mobility on the evolution of cooperation [52–57], although it relies on an essentially different algorithm. The outline of the latter is as follows: 1) Start with a set of strategies (i.e., cooperation probabilities W) that are initially uniformly distributed in the $[0, 1]$ interval. 2) Calculate a velocity vector for each strategy in the swarm. 3) Update the strategy of each agent, using its previous value and the updated velocity vector. 4) Go to step 2 and repeat until convergence. All the details of this setup are described in the Methods section, while here we

proceed with presenting the main results.

Results

We start by presenting the average level of cooperation, defined as $N^{-1} \sum_i W(i)$ where N is the system size and i runs over all the players in the population, in dependence on the temptation to defect b for different values of ω (for the definition see the Methods section) in Fig. 1. Expectedly, the average level of cooperation decreases as b increases for all ω . However, while for $\omega \rightarrow 0$ the cooperative behavior dies out completely at high values of b , for $\omega \rightarrow 1$ the average level of cooperation hovers comfortably over $1/3$, even when the maximal $b = 2$ limit is reached. For intermediate and low values of b , however, small values of ω may yield overall higher average levels of cooperation. It is thus intriguing to find that the introduced particle swarm optimization in the strategy updating, fine-tuned by means of the parameter ω , can be responsible for the emergence of cooperative behavior across the whole span of defection temptation values, as well as for its dominance at low values of b . More precisely, two regimes can be differentiated. For $b < 1.5$ intermediate and high values of ω are actually detrimental for the evolution of cooperation, while for $b > 1.5$ the higher the ω the higher the stationary level of cooperative behavior. These results make it clear that low ω (e.g., $\omega = 0.01$) strongly support the cooperation level for small b , up to $b \simeq 1.2$, whereas high ω are much better suited for cooperation to evolve under this dynamics in strongly defection-prone environments. At this point we argue that for $\omega \rightarrow 1$, when players imitate their best past actions rather than the best players in the swarm (see Methods for details), the proposed strategy updating rule warrants the most significant benefits to cooperative behavior if looking at the entire range of b values, thus in turn resolving the prisoner's dilemma.

In order to obtain an understanding of these results, we first systematically analyze the impact of ω on the final distribution of strategies in the whole population for various values of b , as depicted in Fig. 2. Note that for $\omega = 0.01$ the distribution of strategies is very monotonous, while for $\omega = 0.99$ much more diversity is inferable. Both observations are virtually independent of b . Since the parameter $\omega \in [0, 1]$ determines the tendency of every player to either adopt the most profitable strategy in its past actions ($\omega \rightarrow 1$) or the strategy of the most successful player in its neighborhood ($\omega \rightarrow 0$), these results can be understood very well. In particular, for $\omega = 0.01$ individuals are strongly inclined to imitate the best-performing strategies in the swarm, irrespective of their personal experience in the past. This narrow-sightedness inevitably results in strongly polarized distributions, as only either pure cooperators or pure defectors are the ones most likely to have the overall highest payoffs. Note that this is because the payoffs are directly scaled by W (see Methods). Conversely, for $\omega = 0.99$ the situation is very different since players will focus on their own past actions and learn from them in order to arrive at the best possible strategy. This has the advantage that, unlike for $\omega = 0.01$, here only the immediate neighborhood is explicitly taken into account. For high values of b local considerations are obviously much more important than for low values of b . In the latter case, the nearest neighbors can much easily be neglected since the environment on its own is not strongly favorable for defectors, and hence cooperators can prevail even if overlooking the detailed distribution of strategies in their immediate neighborhood. An additional advantage of small ω , however, is that by focusing only (or predominantly) on the best-performing players in the swarm, the average level of cooperativeness can be maximized more efficiently (as evidenced by results presented in Fig. 1). But if the temptation to defect is strong the strictly local considerations are much more important, as proper adaptation is then crucial for cooperators to survive. Accordingly, for high values of b higher ω yield better results (higher average level of cooperation) by exploiting effectively the whole array of available strategies to respond properly (*locally* properly) to invading defectors. At low values of b , however, these locally optimal adaptations (warranted by $\omega \rightarrow 1$) might be less effective than the more globally inspired actions (warranted by $\omega \rightarrow 0$).

These conclusions can be corroborated further by examining characteristic snapshots of strategy and velocity distributions for key combinations of b and ω , as presented in Figs. 3 and 4. Focusing first on the distribution of strategies in Fig. 3, it can be inferred that for $\omega = 0.01$, where only the most successful strategies within the whole swarm can spread rapidly due to the workings of the particle swarm optimization algorithm, the

strategy distribution becomes very monotonous, leading to the isolation of homogeneous groups of players characterized either by $W = 0$ or $W = 1$, respectively. This holds irrespective of b , only that for strong temptations to defect the clusters of strongly cooperative players become rarer. Note that in this parameter region the here studied stochastic strategy prisoner's dilemma game actually becomes strikingly similar to the classical two-strategy spatial prisoner's dilemma game [34,51], where the clustering of cooperators is the main driving force prohibiting the full dominance of defectors. Conversely, for $\omega = 0.99$, where the particle swarm optimization algorithm is driven by the past experience of every individual player (rather than the swarm as a whole), highly heterogeneous kaleidoscopes appear, and it is indeed this diversity that warrants a high level of cooperativeness even by strong temptations to defect. In particular, snapshots in the bottom panel of Fig. 3 indicate that many clusters consist of a small amount of players with a high cooperation level (i.e., W close to 1), surrounded by players with comparatively lower W values. This in turn implies that not the clustering itself is crucial for the sustenance of cooperation, but actually the aggregation of such clusters itself, which enables the players with higher cooperation level to survive the evolutionary process. Note that the high cooperation level within clusters provides surrounding individuals with a safe source of benefits that are sufficient to resist the invasion of predominantly defective (i.e., W close to 0) players. The particle swarm optimization algorithm thus spontaneously generates the diversity needed for cooperation to survive at high b , much by means of the same mechanism that was reported previously for manually introduced heterogeneous states [58]. Of course, players located in the interior of such clusters enjoy the benefits of mutual cooperation and are therefore able to survive despite the constant exploitation by defectors, yet this positive effect is additionally amplified by the diversity and the hierarchical local structures that give additional strength to the cooperative strategy, while at the same time provide no benefits for defectors.

Moreover, by examining the characteristic distributions of velocities presented in Fig. 4, we can obtain further insight with regards to the evolution of the strategies and their adaptation. Note that by means of Eqs. (1) and (2) (see the Methods section), the two quantities are strongly interdependent. For $\omega = 0.01$, even though the snapshots are taken in the stationary state (where the average level of cooperation is stable), the majority of players will have the velocity very different from 0 (although on average over time and space it is virtually zero, thus assuring the stationary state being reached). This indicates that players will constantly try to reach the currently maximal payoff in the swarm, despite the fact that for the majority this will be unattainable. The locally high velocity values also indicate that the evolutionary process at low values of ω is quite violent and fast, with the population therefore unable to cope with high temptations to defect. Conversely, for $\omega = 0.99$ the situation is very different. Here the majority of players will adapt their strategy very slowly to the changing local influences, which yields the velocity profile for every player being very close to zero. These conclusions are valid practically irrespective of b for the two considered values of ω , but the average level of cooperation is in fact very much different. While individually optimal past strategies in the particle swarm optimization algorithm yield a slow but stable and very effective response even to severe defector attacks, population-wide (or swarm-wide) pursuit for extraordinary benefits proves insufficiently effective to sustain cooperative behavior at high b values. The latter approach, however, may be superior at low temptations to defect, where local considerations are not so vital, and where the pursuit of individual benefits can be successful even if driven by globally-inspired fast and bold actions.

Summary

In sum, we have studied the impact of particle swarm optimization on the evolution of cooperation in the stochastic strategy spatial prisoner's dilemma game. The strategy updating was guided by the particle swarm optimization algorithm, using as input the individual memory of every player (i.e., keeping score of the most profitable individual strategy in the past) as well as the knowledge gained by the swarm (i.e., the nearest neighbors) as a whole. By means of extensive simulations, we found that cooperative behavior can prevail in large regions of the parameter space defining the stochastic strategy prisoner's dilemma game, thus effectively leading to the resolution of the dilemma in favor of pro-social behavior. In particular, we have demonstrated that

imitating the most profitable strategy in the swarm may lead to full dominance of cooperation at moderate temptations to defect, while imitating the best individual actions in the past may lead to the survival of cooperative behavior even if the environment is strongly prone to defection. We have also investigated the actual strategy configurations in the population as well as pertaining spatial distributions of strategies and velocities, for which we have found to be closely tied to the setup of the particle swarm optimization algorithm, and in fact instrumental for the understanding of the observed promotion of the evolution of cooperation. We hope that our work will offer new ways of ensuring cooperation in situations constituting a social dilemma, and that it will be an inspiration for future research when considering the very interesting combination of intelligent algorithms and evolutionary games.

Methods

We consider an evolutionary stochastic strategy prisoner's dilemma game on a square lattice, consisting of 100×100 players with nearest-neighbor interactions and periodic boundary conditions. Initially the strategies of all players are drawn randomly from uniformly distributed values of W in the $[0, 1]$ interval, whereby W determines the cooperativeness of each individual (or the willingness to cooperate). While $W = 1$ returns full cooperation and $W = 0$ returns full defection, between $0 < W < 1$ there exists a continuous set of strategies that can be considered either as being predominantly cooperative (if $W > 0.5$) or predominantly defective (if $W < 0.5$), hence constituting a stochastic strategy version of the prisoner's dilemma game.

Players interact pairwise with all their nearest neighbors, thereby receiving payoffs that can be summarized succinctly by the rescaled payoff matrix

$$\begin{array}{c} C \\ D \end{array} \left(\begin{array}{cc} C & D \\ W(i) * W(j) & 0 \\ b * W(j) * (1 - W(i)) & 0 \end{array} \right)$$

where $W(i)$ and $W(j)$ define the level of cooperativeness of players i and j , respectively. This setup entails b as the only free parameter determining the temptation to defect, but it is well-known that the essence of the prisoner's dilemma game is thereby left intact [34].

The stochastic strategy prisoner's dilemma game is iterated forward in time using a synchronous Monte Carlo updating scheme. First, each player accumulates its payoff by playing the game with all four of its nearest neighbors. Subsequently, players have to decide what strategy they will adopt in the next round (i.e., what will their new $W(i)$ be), which we here determine by means of the particle swarm optimization algorithm. Its implementation is simple and intuitive, as follows. Initially, at time step $n = 0$, all players are assigned the same velocity $V_{i,n} = 0$. For each following n , the velocity vector $V_{i,n}$ of every player i is updated according to

$$V_{i,n+1} = V_{i,n} + \omega[W(i, h) - W(i, n)] + (1 - \omega)[W(\star, n) - W(i, n)], \quad (1)$$

and the strategy follows directly as

$$W(i, n+1) = W(i, n) + V_{i,n+1}, \quad (2)$$

where in Eq. (1) $W(i, h)$ is the most profitable strategy of player i in all its past actions, whereas $W(\star, n)$ is the best performing strategy in the swarm (here considered to be composed of the four nearest neighbors). The parameter $\omega \in [0, 1]$ determines the tendency of every player to either adopt the most profitable strategy in its past actions or the current strategy of the most successful player within the swarm. In particular, $\omega = 1$ implies that the player will definitely imitate its past best action, i.e., the strategy that in the past yielded the highest payoff. On the other hand, $\omega = 0$ implies that the player will copy the strategy of the currently best performing player in its neighborhood. Intermediate values of ω interpolate linearly between these two extremes. Besides the temptation to defect b , ω is here considered as the second crucial system parameter.

References

1. Hofbauer J, Sigmund K (1998) *Evolutionary Games and Population Dynamics*. Cambridge: Cambridge Univ. Press.
2. Nowak MA (2006) *Evolutionary Dynamics*. Cambridge, MA: Harvard University Press.
3. Sigmund K (2010) *The Calculus of Selfishness*. Princeton, MA: Princeton University Press.
4. Nowak MA (2006) Five rules for the evolution of cooperation. *Science* 314: 1560–1563.
5. Hardin G (1968) The tragedy of the commons. *Science* 162: 1243–1248.
6. Axelrod R (1984) *The Evolution of Cooperation*. New York: Basic Books.
7. Doebeli M, Hauert C (2005) Models of cooperation based on prisoner's dilemma and snowdrift game. *Ecol Lett* 8: 748–766.
8. Szabó G, Fáth G (2007) Evolutionary games on graphs. *Phys Rep* 446: 97–216.
9. Roca CP, Cuesta JA, Sánchez A (2009) Evolutionary game theory: Temporal and spatial effects beyond replicator dynamics. *Phys Life Rev* 6: 208–249.
10. Perc M, Szolnoki A (2010) Coevolutionary games – a mini review. *BioSystems* 99: 109–125.
11. Santos FC, Pacheco JM (2005) Scale-free networks provide a unifying framework for the emergence of cooperation. *Phys Rev Lett* 95: 098104.
12. Szabó G, Vukov J, Szolnoki A (2005) Phase diagrams for an evolutionary prisoner's dilemma game on two-dimensional lattices. *Phys Rev E* 72: 047107.
13. Perc M (2006) Coherence resonance in spatial prisoner's dilemma game. *New J Phys* 8: 22.
14. Tomassini M, Luthi L, Pestelacci E (2007) Social dilemmas and cooperation in complex networks. *Int J Mod Phys C* 18: 1173–1185.
15. Gómez-Gardeñes J, Campillo M, Moreno Y, Floría LM (2007) Dynamical organization of cooperation in complex networks. *Phys Rev Lett* 98: 108103.
16. Szolnoki A, Szabó G (2007) Cooperation enhanced by inhomogeneous activity of teaching for evolutionary prisoner's dilemma games. *EPL* 77: 30004.
17. Chen XJ, Fu F, Wang L (2008) Interaction stochasticity supports cooperation in spatial prisoner's dilemma. *Phys Rev E* 78: 051120.
18. Poncela J, Gómez-Gardeñes J, Floría LM, Sánchez A, Moreno Y (2008) Complex cooperative networks from evolutionary preferential attachment. *PLoS ONE* 3: e2449.
19. Pestelacci E, Tomassini M, Luthi L (2008) Evolution of cooperation and coordination in a dynamically networked society. *Biol Theor* 3: 139–153.
20. Fu F, Wu T, Wang L (2009) Partner switching stabilizes cooperation in coevolutionary prisoner's dilemma. *Phys Rev E* 79: 036101.
21. Poncela J, Gómez-Gardeñes J, Floría LM, Moreno Y, Sánchez A (2009) Cooperative scale-free networks despite the presence of defector hubs. *EPL* 88: 38003.

22. Van Segbroeck S, Santos FC, Lenaerts T, Pacheco JM (2009) Reacting differently to adverse ties promotes cooperation in social networks. *Phys Rev Lett* 102: 058105.
23. Szabó G, Szolnoki A, Vukov J (2009) Selection of dynamical rules in spatial prisoner's dilemma games. *EPL* 87: 18007.
24. Szabó G, Szolnoki A (2009) Cooperation in spatial prisoner's dilemma with two types of players for increasing number of neighbors. *Phys Rev E* 79: 016106.
25. Peña J, Volken H, Pestelacci E, Tomassini M (2009) Conformity hinders the evolution of cooperation on scale-free networks. *Phys Rev E* 80: 016110.
26. Wu ZX, Holme P (2009) Effects of strategy-migration direction and noise in the evolutionary spatial prisoner's dilemma. *Phys Rev E* 80: 026108.
27. Poncela J, Gómez-Gardeñes J, Traulsen A, Moreno Y (2009) Evolutionary game dynamics in a growing structured population. *New J Phys* 11: 083031.
28. Wu B, Zhou D, Fu F, Luo Q, Wang L, et al. (2010) Evolution of cooperation on stochastic dynamical networks. *PLoS ONE* 5: e11187.
29. Rong Z, Wu ZX, Wang WX (2010) Emergence of cooperation through coevolving time scale in spatial prisoners dilemma. *Phys Rev E* 82: 026101.
30. Cardillo A, Gómez-Gardeñes J, Vilone D, Sánchez A (2010) Coevolution of strategies and update rules in complex prisoner's dilemma networks. *New J Phys* 12: 103034.
31. Szolnoki A, Wang Z, Wang J, Zhu X (2010) Dynamically generated cyclic dominance in spatial prisoner's dilemma games. *Phys Rev E* 82: 036110.
32. Fu F, Nowak MA, Hauert C (2010) Invasion and expansion of cooperators in lattice populations: Prisoners dilemma vs. snowdrift games. *J Theor Biol* 266: 358–366.
33. Perc M, Wang Z (2011) Heterogeneous aspirations promote cooperation in the prisoners dilemma game. *PLoS ONE* 5: e15117.
34. Nowak MA, May RM (1992) Evolutionary games and spatial chaos. *Nature* 359: 826–829.
35. Hauert C, Doebeli M (2004) Spatial structure often inhibits the evolution of cooperation in the snowdrift game. *Nature* 428: 643–646.
36. Abramson G, Kuperman M (2001) Social games in a social network. *Phys Rev E* 63: 030901(R).
37. Zimmermann MG, Eguíluz V, Miguel MS (2004) Coevolution of dynamical states and interactions in dynamic networks. *Phys Rev E* 69: 065102(R).
38. Santos FC, Pacheco JM, Lenaerts T (2006) Evolutionary dynamics of social dilemmas in structured heterogeneous populations. *Proc Natl Acad Sci USA* 103: 3490–3494.
39. Poncela J, Gómez-Gardeñes J, Floría LM, Moreno Y (2007) Robustness of cooperation in the evolutionary prisoner's dilemma on complex systems. *New J Phys* 9: 184.
40. Perc M (2009) Evolution of cooperation on scale-free networks subject to error and attack. *New J Phys* 11: 033027.
41. Hamilton W (1964) The genetical evolution of social behaviour. I. *J Theor Biol* 7: 1–16.

42. Nowak MA, Sigmund K (1998) The dynamics of indirect reciprocity. *J Theor Biol* 191: 561–574.
43. Brandt H, Sigmund K (2005) Indirect reciprocity, image scoring, and moral hazard. *Proc Natl Acad Sci USA* 102: 2666–2570.
44. Brandt H, Sigmund K (2006) The good, the bad and the discriminator - errors in direct and indirect reciprocity. *J Theor Biol* 239: 183–194.
45. Dugatkin LA, Mesterton-Gibbons M (1996) Cooperation among unrelated individuals: reciprocal altruism, by-product mutualism and group selection in fishes. *Biosystems* 37: 19–30.
46. Traulsen A, Nowak MA (2006) Evolution of cooperation by multilevel selection. *Proc Natl Acad Sci USA* 103: 10952–10955.
47. Szolnoki A, Perc M (2009) Emergence of multilevel selection in the prisoner's dilemma game on coevolving random networks. *New J Phys* 11: 093033.
48. Kennedy J, Eberhart R (1995) Particle swarm optimization. In: *IEEE International Conference on Neural Networks*. Piscataway, volume 4, pp. 1942–1948.
49. Eberhart R, Kennedy J (1995) A new optimizer using particle swarm theory. In: *Proc. Sixth Intl. Symp. on Micro Machine and Human Science*. Nagoya, Japan: New York, NY, USA: IEEE, volume 43.
50. Eberhart R, Simpson P, Dobbins R (1996) *Computational intelligence PC tools*. Academic Press Professional, Inc. San Diego, CA, USA.
51. Szabó G, Tóke C (1998) Evolutionary prisoner's dilemma game on a square lattice. *Phys Rev E* 58: 69–73.
52. Vainstein MH, Arenzon JJ (2001) Disordered environments in spatial games. *Phys Rev E* 64: 051905.
53. Vainstein MH, Silva ATC, Arenzon JJ (2007) Does mobility decrease cooperation? *J Theor Biol* 244: 722–728.
54. Meloni S, Buscarino A, Fortuna L, Frasca M, Gomez-Gardeñes J, et al. (2009) Effects of mobility in a population of prisoner's dilemma players. *Phys Rev E* 79: 067101.
55. Helbing D, Yu W (2009) The outbreak of cooperation among success-driven individuals under noisy conditions. *Proc Natl Acad Sci USA* 106: 3680–3685.
56. Droz M, Szwabinski J, Szabó G (2009) Motion of influential players can support cooperation in prisoner's dilemma. *Eur Phys J B* 71: 579–585.
57. Jiang LL, Wang WX, Lai YC, Wang BH (2010) Role of adaptive migration in promoting cooperation in spatial games. *Phys Rev E* 81: 036108.
58. Perc M, Szolnoki A (2008) Social diversity and promotion of cooperation in the spatial prisoner's dilemma game. *Phys Rev E* 77: 011904.

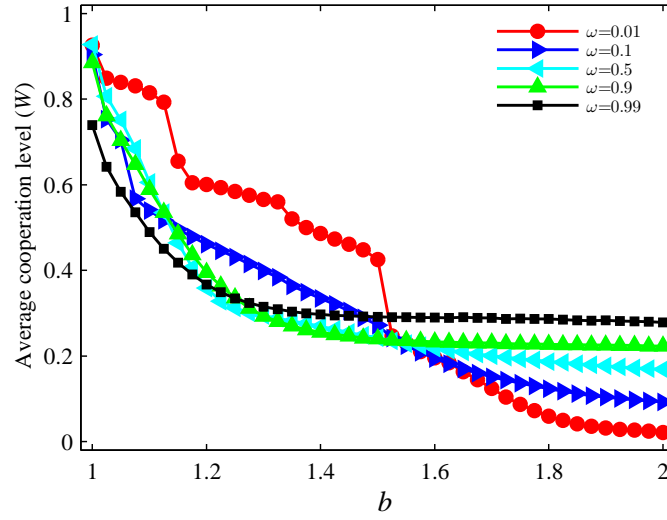


Figure 1. Average level of cooperation in dependence on b for different values of ω . It can be observed that while imitating the best performing player in the swarm ($\omega \rightarrow 0$) might be beneficial at low temptations to defect, imitating personal success ($\omega \rightarrow 1$) is definitively better for the evolution of cooperation in strongly defection-prone environments. Each data point is an average of the final outcome (stationary state) of the game over 100 independent realizations. Lines connecting the symbols are just to guide the eye.

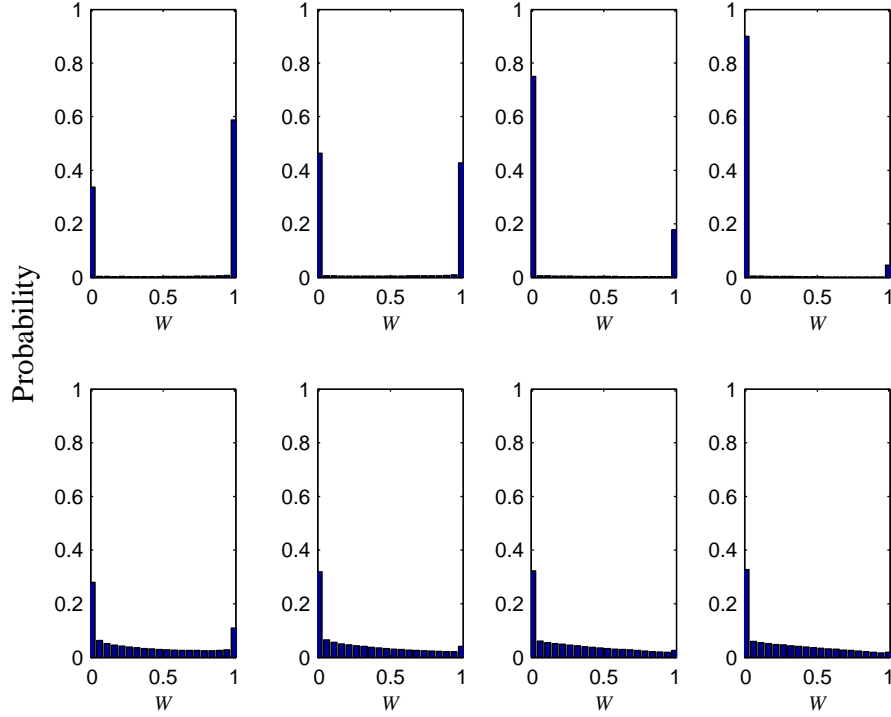


Figure 2. Distribution of strategies in the whole population, as obtained for different combinations of b and ω . It can be observed that for $\omega = 0.01$ the nature of the stochastic strategy prisoner's dilemma game is essentially completely overridden by the selfish drive of players to reach the highest current payoffs in the swarm, in turn virtually completely transforming the game to its two-strategy [only $W = 0$ (full defection) or $W = 1$ (full cooperation) strategies are present in the population] version. Conversely, for $\omega = 0.99$ the full spectrum of available strategies is exploited to arrive at the final stationary state. Note that the horizontal axis displays the willingness to cooperate W (defining the strategy of every player), while the vertical axis depicts the probability that this strategy is present in the population. Depicted results are averages of the final outcome (stationary state) over 100 independent realizations.

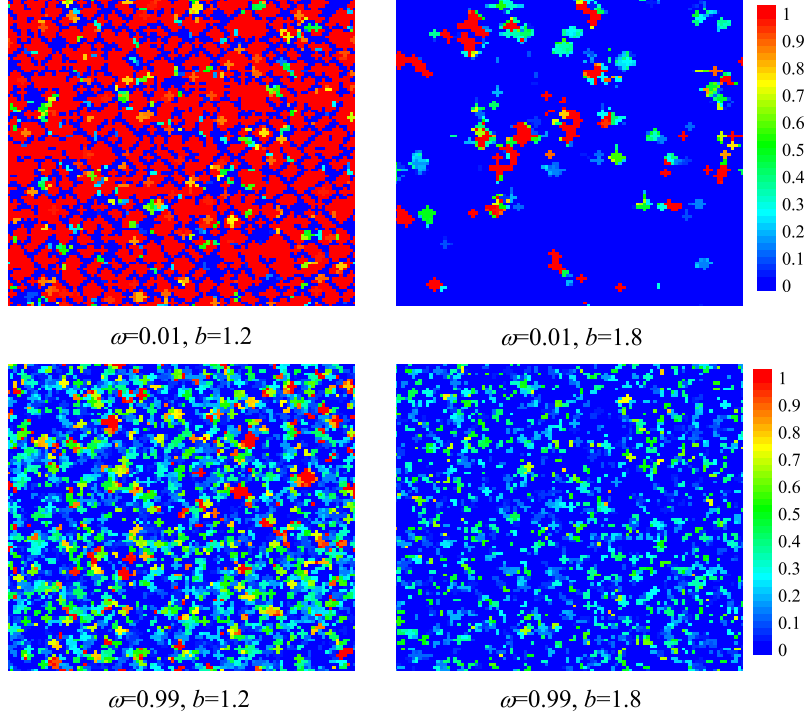


Figure 3. Characteristic spatial distributions of strategies, as obtained for different combinations of b and ω . As concluded from results depicted in Fig. 2, for low values of ω only the two “extreme” strategies (with rare exceptions) are adopted, while for high values of ω the whole array of available strategies comes into play. Moreover, it is interesting to observe that values of $\omega \rightarrow 0$ yield the well-known clustering of cooperators [34] on the square lattice, while the snapshots for $\omega = 0.99$ seem to have these feature somewhat less pronounced, although still clearly inferable (note that the distinction of clusters is somewhat difficult due to the continuous array of possible strategies). This suggests that, besides the clustering of cooperators, additional mechanisms may underlie the survival of cooperators at high temptations to defect and $\omega \rightarrow 1$ within the present setup. The color encoding, as depicted right, indicates the values of W for each individual player.

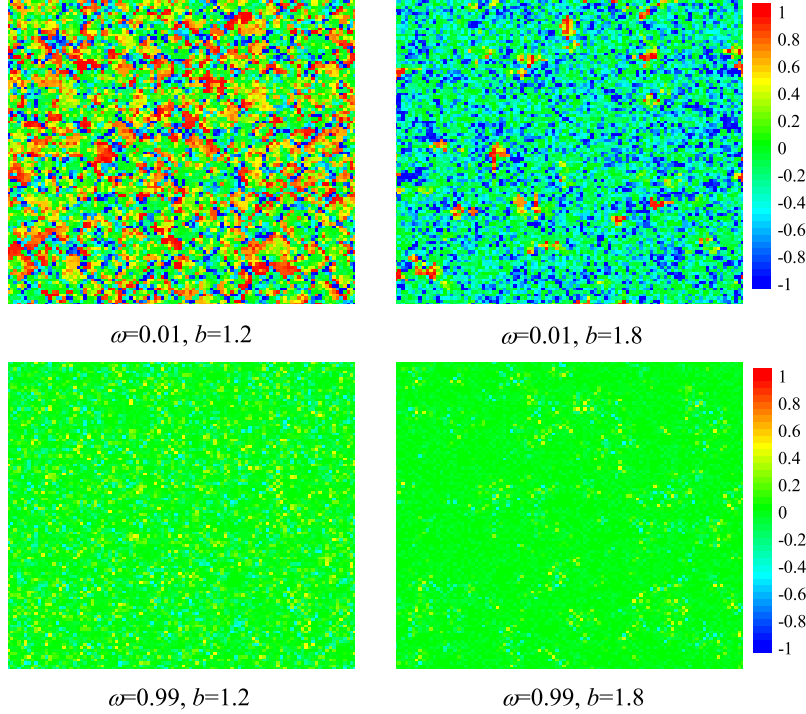


Figure 4. Characteristic spatial distributions of velocities, as obtained for different combinations of b and ω . Top row depicts results for $\omega = 0.01$, while bottom row features results for $\omega = 0.99$. Irrespective of b , it can be observed that for $\omega = 0.99$ the whole population essentially becomes a swarm in that the velocities of all players are much the same and close to zero. The fact that the prevailing velocity is close to zero simply reflects that the stationary state has been reached by means of adaptive, locally-inspired and slow strategy changes (which are, however, very effective even if the temptations to defect are strong). For $\omega = 0.01$, however, only isolated clusters can be considered to act as swarms, while the majority of players cannot be associated with any kind of group dynamics and is simply caught in the futile pursuit for the highest, yet for the majority unattainable, payoffs. These results indicate that swarming is an important agonist that promotes cooperation at high temptations to defect (see results presented in Fig. 1). The color encoding, as depicted right, indicates the values of $V_{i,n}$ for each individual player, where n was chosen sufficiently large such that the stationary state of the game has been reached. Importantly, we note that for $\omega = 0.01$ the stationary state has in fact been reached, although at a given instance in time the average velocity in the population might be different from zero.